DOCUMENT RESUME

ED 234 631 FL 013 952

AUTHOR Wald, Benji
TITLE Report on the Study of Limited Language Proficiency.
INSTITUTION National Center for Bilingual Research, Los Alamitos, Calif.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
REPORT NO NCBR-01-80
PUB DATE Sep 80
NOTE 37p.
AVAILABLE FROM National Center for Bilingual Research, 4665 Lampson Ave., Los Alamitos, CA 90720.
PUB TYPE Viewpoints (120)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Bilingualism; Elementary Education; *Language Proficiency; *Language Tests; Test Bias; Test Construction; *Testing Problems; Test Interpretation

ABSTRACT

The concepts of language proficiency, limited language proficiency, and comparably limited language proficiency are examined. Studies critical of the instruments used for making assessments according to these concepts are reviewed. The effects of these concepts on proposed legislation involving bilingual education are discussed. It is suggested that linguistic proficiency test results may be invalid because of problems in test administration and design or problems inherent in the child. A systematic attempt to describe the language behavior of comparably limited proficient children is suggested. Finally, it is hypothesized that (1) most language proficiency instruments are not an accurate indicator of actual language proficiency and school achievement; (2) studies of speech behavior would provide a more accurate diagnostic tool; (3) the comparably limited category is largely an artifact of inaccurate instrumentation and testing bias; and (4) school achievement does not directly correspond to tested or actual language proficiency. (Author/RW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED234631

REPORT ON THE STUDY OF LIMITED LANGUAGE PROFICIENCY

Benji Wald

September, 1980

01-80

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

NCBR

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

FL013952

national center for bilingual research, 4665 lampson avenue, los alamitos, CA 90720, telephone (213) 598-0481

The material in this publication was prepared under Cooperative Agreement 00-CA-80-0001 from the National Institute of Education, Department of Education. Its contents do not necessarily reflect the views of the National Institute of Education or of any other agency of the United States government, and no official endorsement should be inferred.

Table of Contents

	Page
ABSTRACT	iii
Introduction	1
What is Language Proficiency?	1
What Do Language Proficiency Tests Measure?	3
How Large is the "Comparably Limited" Population?	7
What Does "Limited Language Proficiency" Mean?	11
A) Problems inherent in the child	12
B) Problems in test administration	16
Problem area #1: Resistance/unfamiliarity on the part of the child	16
Problem area #2: Mechanical and other errors on the part of the tester	17
C) Problems in test design	18
Problem area #1: Irrelevant or wrong questions or tasks for assessment	19
Problem area #2: Built-in discriminatory factors against nonstandard language features	21
Problem area #3: Uncontrollable fluctuations in performance on the part of the testee, giving a false impression of stable competence	24
Issues and Questions	24
Summary	26
Hypotheses	26
REFERENCES	28

ABSTRACT

This report discusses the notion of "language proficiency," "limited language proficiency," and ultimately "comparably limited (language proficiency)" as concepts in need of further definition and clarification. Studies critical of the test instruments used for making assessments according to these concepts are presented. The author goes on to discuss the specific effect these concepts have had on proposed legislation affecting bilingual education. Then, some sociolinguistic considerations are suggested for why test results may be inaccurate and/or invalid. Finally, five major points are summarized having a bearing on continued research in this area, and four hypotheses are proposed concerning the relationship of conventional language proficiency testing, actual speech behavior in naturalistic situations and school achievement.

REPORT ON THE STUDY OF LIMITED LANGUAGE PROFICIENCY

Benji Wald

Introduction

The term "semilingualism" (further discussed below) has not received wide currency in the discussion of American educational problems relevant to bilingual education, although the underlying concept is much evident in such documents as the Federal Register under the label "comparably limited," which will be herein referred to as "assessed comparably limited," or "alleged comparably limited" (hereafter ACL) on the basis of the failure of children of this group to attain a "proficient" status in any tested language according to one or another of a number of currently competing language proficiency assessment instruments.

Both the general issue of language proficiency assessment and the more particular issue of whether the category "comparably limited" is a valid one need rigorous study which goes beyond current work. In terms of research strategy, the two issues conflate. There is reason to believe that the category "comparably limited" has not been demonstrated to be valid for the majority of the children so-labled; we hypothesize, then, that defects other than actual language proficiency are responsible for the resultant assessments, most notably in the administration, design, and underlying assumptions of the tests. Under the hypothesis presented here, at least part of the responsibility for the lack of validity of the category "comparably limited," and of language proficiency tests in general, is the lack of a sufficiently developed notion of what language proficiency is and how it can be measured in a way relevant to projected school achievement.

What is Language Proficiency?

Underlying conventional test assessment is a concept of language proficiency based on the common sense notion that people have different degrees of competence (knowledge in the Chomskyan sense) in various

languages, ranging from "none" in most languages to something indeterminate in at least one language. This "indeterminate" proficiency in a particular language has no upper limit, but it is expected that in the overwhelming number of cases it is based on sufficient competence for most purposes, and most importantly for receiving meaningful education by means of that language (although it would probably not be appropriate to require writing or critiquing Shakespearean drama as a measure for facility in a language). It is the necessity of having the linguistic skills to benefit from U.S. schooling that is the reason behind tests of language proficiency. The major impetus for their use has been the 1974 Supreme Court Lau decision, which precipated the need for such tests to facilitate an equitable "meaningful education of non-English-proficient students in the public eduction system.

If the notion of "comparably limited" is taken seriously, the common sense notion of linguistic competence, i.e., the ability to function well in daily verbal interaction, applies to only "some" people. It would mean that monolinguals have sufficient competence for most purposes in one language, while others, particularly some bilinguals (with the broadest possible understanding of the term "bilingual"), do not have sufficient competence in any language to determine the language in which they may best receive a meaningful education. This will be further discussed below. For the moment it is necessar? to pursue the common sense notion of language proficiency in order to see what it is based on.

Linguistic competence involves both speaking and comprehension, of which only speaking is directly observable for assessment. Comprehension is inferred from the effect of language on the behavior of a listener. (In natural situations this is not always possible since, for example, comprehension may only be evident in the fulfilling of a directive that to take place at a later time, e.g. "drop me a line when you get to London"; and, from the point of view of an interlocutor,

comprehension is often taken for granted so that simple back-channeling cues like "uhunh" or an assentive nod are interpreted as understanding, whereas this may not necessarily be the case.) While it is beyond doubt that a certain degree of language proficiency is necessary in order for a person to benefit from instruction in that language, it is not obvious how much proficiency is needed or that past a certain point, the more proficiency the better the achievement.

What do language proficiency tests measure?

Language proficiency tests are based on various assumptions about what to measure and how to interpret results. Quite generally these instruments emphasize structural features of performance and some, in addition, the following of non-functional directions. It has been pointed out by some observers that there is a lack of match between current concepts of communicative competence, which focus on ability to function interactionally in a language, and test procedures which evaluate aspects of linguistic structure divorced from any social context (e.g., Galang, 1980). That is, language proficiency develops and is used in everyday social contexts, where functional and pragmatic features are interwoven with operation of the formal (structural) features of language. In contrast, most tests generally denude language structure of its normal functions and place speakers in a highly unusual situation whose effects on linguistic behavior are very problematic, although largely ignored, in using the instrument to assess the proficiency of the subject.

It is far from certain that the most widely used tests of language proficiency accurately distinguish those who are proficient in a particular language from those who are not. It is possible, as hypothesized below, that those who fail to be recorded as proficient include a significant number who are in fact proficient; and it is also possible that those who are registered as proficient (or "functional") include a number who are actually not or, at least, would benefit more

by partial instruction in another language (the home language). (The second point above was brought up by Luis Laosa at the New York City hearings for the proposed Federal Regulations guiding funding for bilingual education in September, 1980.)

Moreover, the two possibilities that both proficient and non-proficient pupils are overestimated by the same instruments needn't be contradictory, since the reasons for overestimation may be different in both cases. From a strictly linguistic point of view, Dietrich & Freeman (1979) discuss a large number of examples from tests which would tend to underestimate a speaker's oral proficiency, mainly due to scoring procedures for syntactic complexity. However, they also point out defects in various tests where faulty assumptions or design may lead to overestimation of proficiency, e.g., some tests generalize "overall" proficiency from knowledge of textually infrequent words (in elementary readers circa 1952), some tasks involve selecting a picture from a set of dissimilar pictures, which can be done on the basis of knowledge of one or two lexical items rather than complete comprehension of the syntactic form of the oral instruction.

Several studies indicate that the non-functionality of the usual language assessment instruments may result in underestimation of actual language proficiency under the hypothesis that situation is a major determining factor of language output (cf. Labov, 1972). For example, work at the Center for Reading has indicated that Black children's output of well-formed utterances and number of words and sentences increases in describing pictures from the Peabody test when the subject is placed at a distance from the tester so that it is no longer obvious to the subject that the test can also see the picture. Removing the subject and tester from close physical proximity to each other and the tester from access to the picture creates an apparent functional reason for greater verbality on the subject's part (Guthrie & Steffensen, 1980).

Ongoing research at the Southwest Educational Development Laboratory involving a longitudinal study of the oral language development of Texas bilingual children (Spanish-English) in which kindergarten and first grade children are periodically recorded in the domains of classroom (to teacher), playground (to peer) and home (to family member), reports that in both quantity and quality displayed language proficiency differs systematically such that greatest proficiency is generally displayed in the playground (Mace-Matluck, 1980, esp. p. 7). Although in that study only structural characteristics of language are being investigated, the implications of the effect of function on language are clear.

Besides the effect of situation on displayed language proficiency, there have been several studies which have thrown doubt on the intrinsic validity of some of the test instruments as indicators of language proficiency at all or as relevant indicators of scholastic achievement. Extensive criticism of the linguistic assumptions of various tests or their particular subparts is found in Dietrich and Freeman (1979). Rosansky (1978) focuses extensive criticism on the BSM (Bilingual Syntax Measure) for various measures of validity which apply wholly or in part to other tests as well. In addition, several normative studies have been injurious to the claims made by some test developers for the validity of their instruments. For example, a study in the Houston area (Gillmore & Dickerson, 1979), in which pairs of five language proficiency test instruments were compared for equivalence (BSM; LAS--Language Assessment Sacles; PAL--Primary Acquisition of Language; SPLIT--Schutt Primary Language Indicator Test; and BINL--Basic Inventory of Natural Language, the last of which had three successive versions for calculating raw score and consequent proficiency category on the basis of the same data), found that there was a low to moderate correlation at best between the classifications obtained from any two instruments, perhaps not surprisingly since the tests disagree as to what should be tested to obtain a proficiency measure. Thus, the distinction between limited and non-limited proficiency (in ?nglish) depends on which test

is used. In addition, there was never a high correlation between any of the proficiency tests and any of the achievement subtests used by the Texas Educational Agency for exit criteria from a bilingual program. Indeed, in the case of the BINL, the correlation between language proficiency assessment and achievement score was negative in most cases.

Another study from the Chicago area (Perlman & Rice, 1979) focused on the LAS exclusively and found that it identified varying proportions of monolingual English speaking students with slightly higher than average reading scores (according to the city-wide mean) as non- or limited-English-proficient, such that the proportion of students assessed as proficient increases with age. This is to be noted, for the LAS claims to be age-independent up to the age of 11, at which point the ceiling for proficiency rises and the study again shows an increase in the number of assessed limited-English-proficiency children. The LAS identified 77.8% of the monolingual English speakers as limited at age 5. This decreases continually to 12.5% at age 11. At age 12, where LAS is renormed, limited proficiency again increased to 40% and then gradually decreased to 25% at age 14.

While the above-cited studies have been concerned only with English, there is evidence that the same defects apply to other language tests, of which, at present, only instruments in Spanish are sufficiently developed to be examined in similar ways. One such study (Jackson, 1980) from the Texas bilingual children oral language development study, indicates that there is a vast lack of agreement between LAS as an assessment instrument for Spanish (mean 2.59) and teachers' ratings of the same kindergarten and first grade children (mean 4.17). In that study the overestimation of non-proficiency when compared to teacher ratings also applies to English, although at much lower level (mean LAS score, 2.10; mean teacher rating, 2.89) than for Spanish. Consequently, there is some reason to be suspicious of the validity of a "comparably limited" category.

Assuming that these problems ultimately derive from confusion and lack of clarity about what language proficiency in fact consists of, the most crucial population against which to consider this issue are those children who are assessed to be comparably limited in two languages. In the following sections, the language policy and theoretical considerations concerning this population are presented.

How Large Is the "Comparably Limited" Population?

There is a large but imprecisely known group of LEP (limited English proficiency) students in the American school system who are currently classified as comparably limited (henceforth CL), that is, also limited in their primary language. Strictly speaking, classification as LEP and/or the subcategory CL is based on achieved performance on a language proficiency assessment instrument, of which a large number exist, e.g., four are in current use in the California school system, BINL, BSM, LAB, LAS, one of which as been selected according to the individual school district. The actual implications of LEP and CL classification in terms of actual language proficiency, everyday language use in and out of school, effect on school achievement, and on teacher's perceptions and expectations of such labeled students, have not been carefully studied and are not known. However, these classifications have been of great importance to the proposed legislative guidelines controlling minimal requirements for bilingual education. Consequently, it is of extreme importance to discover the realities underlying these labels and their effects in the school system.

The size of the assessed comparably limited (ACL) population among LEP students is a matter of some controversy. No systematic study has been done in this area. Burt and Dulay (1980) report variable numbers of ACL depending on district in a study of several school districts in California, ranging from a high of 94% of all LEP students sampled in Santa Clara (out of a total of 166 children in grades K-8 using the BSM

instrument) to a low of 10% of all LEP students sampled (576 in grades K-2 using the BSM) in San Diego. Methods of sampling have not yet been made clear to us. It is difficult, therefore, to tell how representative these data are. However, they follow a pattern such that non-proficiency in Spanish increases as districts increase in distance from the Mexico/U.S. border area, a pattern which is repeated in other studies (cf. Laosa, 1977). De Avila and Duncan (1979) report an ACL proportion of 26.5% of their total sample (56 out of a 191 sample) according to the LAS instrument. This they divide into two categories, 1) limited bilingual (20.6% of total), and 2) late language learner (5.9% of total), the difference being that the late learners fail to score a 3 or above in either language (of English and Spanish).

The distribution of ACLs differs according to research site. Given the four sites over which the total sample was distributed, the Puerto Rican-Americans (at an unspecified site in the Northeast, but not New York City) had the highest incidence at 45.6% and the Cuban-Americans (at a site in Miami) had the lowest incidence at 9.3%. Of the two Mexican-American groups studied, the urban group (at a site near San Francisco) had an incidence of 29.9%, while the rural group (At a site near El Paso and, thus, near the Mexican border) had a lower incidence of 15.7%. From their figures one can discern a pattern such that the population of late learners at any site grows as the number of limited bilinguals increases, as if the causes underlying both categories are the same. For example, the Puerto Rican group, with the greatest number of limited bilinguals, also had the greatest number of late learners (15.6%), while the Cuban group had no late learners.

In discussion, the authors refer to a previous study (Linguametrics, 1977) indicating that ACL incidence ranged from 20% in suburban studies to "as high as 36% in inner city urban schools" (op. cit. p. 28). It is not clear from the report how the sample was actually selected, prohibiting an understanding of its representativeness to larger populations. In the Texas longitudinal

study using the LAS instrument, 76% of the kindergarten and first graders fit the ACL category (as either "limited bilinguals" or "late language learners," the latter being the majority).

In the following section possible reasons for the differential distribution of ACL across different social groups can be discussed, under the assumption that such figures are to some extent representative. Up to this point, the object has been to present available evidence that the category is of sufficient size to merit special interest on its own. We will now proceed to examine how this sizeable population is affected by proposed educational policies.

To underscore the importance of the issue of the reality underlying the ACL category as it applies to the U.S. educational policy, it is worth citing at length from the Federal Register of August 5, 198, which discusses the proposed rules for bilingual education services (pp. 52056-7):

> (a). . .Limited proficiency students who are clearly English superior need not be given specially tailored services. These students rely primarily on English. It is their strongest language. The type of help they may need in improving their English skills is similar to that needed by monolingual speakers of English who have relatively weak language skills. The proposed rules, therefore, simply mandate that English-superior students have the same access to compensatory help as other students.

Pertinent to this citation, a comment is necessary. Theoretically, this proposal seems eminently reasonable. In practice, however, further specification is necessary concerning what "clearly English-superior" means. Given that a student is LEP to begin with, how much lower does his/her score have to be in the primary language in order to rate as "clearly English-superior"? And, to take up a theme which will be further discussed and exemplified later, how sensitive and fine-tuned does the instrument for language proficiency assessment

need to be in order for differences in scores in two languages to be significant (i.e., clear). This is far from an academic question since in some cases test scores fluctuate greatly from one testing session to another, changing assessment of which is the stronger language.

To continue with the citation:

> (b). . . The remaining students--those who are comparably limited and those who are clearly primary-language superior--must receive instruction designed to develop full proficiency in English. . .

A similar objection must be raised to the word "clearly" in this passage. In effect, although we are not challenging at this point the provision for several subtypes of CL students, we are insisting that careful research is necessary to insure that these distinctions are accurately made, and not on the basis of instruments that are not clearly defined nor equal to the task.

The next passage defines the only group of LEP students who will unequivocably get any support for bilingual education:

> (c). . . Because students with weak English skills may fall permanently behind in other required subjects while they are learning English, the proposed rules require that primary-language-superior students receive instruction through both languages in required subjects while the students are learning English. Instruction in English may be increased as the student's command of the English language increases.

At this point we come to the most important segment of the citation with regard to the treatment of CL students:

> Two alternatives are presented in the text of the proposed rules to stimulate comment on whether <u>comparably limited</u> students should also receive this instruction. These alternatives are presented because experts disagree about which placement is best for comparably limited students. Some educators argue strongly that students who are comparably limited in two languages will have a difficult time in a monolingual, English-speaking class. They also

> argue that the skills that limited-English-proficient students possess in their own language are not necessarily the same as those possessed in English. For example, vocabulary or grammar skills may be different in each language. Others object to this assumption and argue that these students will do better in the long run if their English skills are sharpened by instruction offered exclusively through English.

For the moment we must suspend discussion of who these experts and educators are, and to what extent their arguments are based on speculative or empirical considerations. It is clear from the passage that nothing has been resolved about this problem. No research has been done with sufficient appeal to legislators to resolve it.

Until research is done in this area, a large number of students fitting into the CL category are in danger of being misclassified and deprived of the meaningful education to which they are entitled under Title VI of the Civil Rights Act of 1964 and the Supreme Court Lau decision of 1974.

The following sections deal with the theoretical considerations and initial plan of work for research on the problem of CL students.

What Does "Limited Language Proficiency" Mean?

Despite its long history and widespread use, the notion of "limited language proficiency" is unclear and ambiguous. In the context of current bilingual education and the Lau guidelines, "limited language proficiency" refers to failure on the part of the individual student to attain a certain cutoff point on a given language proficiency test instrument. Therefore, one point of contention can be--and is--where the cutoff point separating limited from non-limited language proficiency should be. In addition, given that different tests are given in different districts to assess language proficiency, to what extent do the cutoff points agree across tests? If they do not, then a student may be classified as limited on one test and

non-limited on another. Which test is more accurate? A study of this type for the four tests used in the state of California has been commissioned by California's Office of Bilingual/Bicultural Education and should be available shortly. This information will be reviewed when available. The Gillmore and Dickerson (1979) study in Texas has already been mentioned above and does not bode well for comparability across tests.

Another question to be raised is, what are the sociolinguistic implications of the classification "limited language proficiency" on tests for actual language proficiency? There are various possibilities, none of which can be dismissed a priori:

A) Problems inherent in the child.

B) Problems in test administration.

C) Problems in test design.

A) Problems inherent in the child. This is the possibility that the child actually lacks necessary proficiency in the relevantly tested language. The meaning of the word "relevantly" in the above sentence will be clarified when the difference between standard and nonstandard varieties of a language is discussed. For the moment we need only note that the purpose of language proficiency tests is to determine whether or not the child has sufficient knowledge of the particular language tested to receive educational instruction in that language. To the extent that possibilities B) and C) above are responsible for scores indicating lack of proficiency, the tests fail to correctly and adequately assess the child's proficiency in the tested language. Before discussing these possibilities, we will begin with considerations under A).

The most serious cases of limited language proficiency apply to monolingual as well as bilingual children. These are the cases in which limited language proficiency is due to some biologically based

impairment, e.g. autism, a neurological impairment, some other physiologically based form of mental retardation or schizophrenic condition. In these cases a child will fail to acquire language despite a supportive environment and will inevitably test on any proficiency instrument as limited and, furthermore, will need some special form of education for the handicapped rather than necessarily bilingual or intensive English training. Too large a percentage of cases of CL proficiency is reported to make this possibility plausible for the majority of cases involving bilinguals (cf. Attanasi et al., 1978; Burt & Dulay, 1980b; Skutnabb-Kangas & Toukomaa, 1976; de Avila & Duncan, 1980; Mace-Matluck, 1980).

An even less plausible possibility is that the CL is of normal capability to acquire language but comes from an impoverished background in which s/he was given insufficient exposure to any language. This idea is embodied in the "verbal deprivation" theory popular among educational psychologists in the late 1960s but largely discredited due to massive sociolinguistic and dialectological research on nonstandard languages of non-middle class populations (among these studies Labov, 1968 and 1972, are classic). The only valid cases where "verbal deprivation" is applicable are extremely rare cases of antisocial upbringing, such as Genie and other "wolfchildren" who fail to acquire language because of extreme deprivation caused by virtually total isolation from human contact. Such cases are immediately recognizable because victims of this treatment fail to acquire all other social behaviors in addition to any form of language. Such cases are not an issue here, but indicate where a serious form of the "verbal deprivation" theory has some validity.

Sociolinguistic and psycholinguistic research over the last two decades has amply demonstrated that the vast majority of children in any linguistic community acquire the language of that community regardless of caregiver practice in overtly teaching the child to speak or even speaking to him/her. However, the <u>variety</u> of language spoken

in the community may be quite different from the standard language of the schools serving that community. To the extent that this problem is reflected in low scores on a language proficiency instrument, the problem is one of test design and not the child. This will be further discussed under the appropriate heading (C).

While "verbal deprivation" has been massively discredited for children of the U.S. Black community due to the research of Labov and others, it has been applied in a somewhat revised form to children in bilingual communities. The general form of the argument is that, in some sense, many children who acquire language in a bilingual community do not get adequate exposure to either language, perhaps because they are unable to separate out which language is which, or because they are exposed to intrasentential codeswitching and other forms of language mixing in which two languages are inextricably mixed in the same context. Some scholars point to such codeswitching as the result of the inability to control either language adequately (cf. Gonzales, 1977). This matter needs more exploration, since scholars are not at all clear on what codeswitching is or on how many types there are, and imply that there may be inherent disadvantages to acquiring language in a bilingual community, a claim which would be patently false if applied to any particular monolingual community. One piece of contradictory evidence is Poplack, 1979, who presents extensive evidence for a Puerto Rican bilingual community in New York City that, contrary to the "codeswitching as limited proficiency in either language" theory, codeswitching in that community is a style of speech sensitive to social situation where intrasentential codeswitching is typical of speakers highly skilled in both languages rather than of those who are limited in one or both languages.

Given what is currently known about language and its acquisition, it is highly unlikely that the normal child (of bilingual as well as monolingual background) is limited in all of the languages s/he has been exposed to. If s/he is limited in one language, it is likely that

s/he is relevantly proficient in another. Here "relevantly proficient" means proficient for age-level in the language(s) used as the primary vehicle(s) of communication in the particular community. This language (or these languages) may be nonstandard, i.e., not recognized as independent languages (as opposed to dialects of a language) by the institutions serving the community (e.g., the school system). It may not be written. It may even not be overtly valued by adult users themselves due to insecurity caused by pressures to conform to official norms. However, in terms of language acquisition, there is no evidence that nonstandard language is any different from a standard one. It is totally false to claim that someone who knows a nonstandard language but not a standard one does not have proficiency in any language. Since it is unlikely that the large number of children classified as CL are less normal in language acquisition ability than their monolingual peers and do not have proficiency in any language, it is unlikely for most of these children that the category CL is a valid one.

We point out here that the category CL is a conceptual offspring of the term "semilingual" popularized by Skutnabb-Kangas and Toukomaa (1975, 1976) in their studies of the educational achievements of the children of Finnish migrant workers in Sweden and still widely used by scholars in Europe and Asia. The authors define a semilingual as "a person who does not know any language properly." It may be the case that "properly" means "standardly" in this context, but this is not clear from the exchange between the authors (1979) and Brent-Palmer (1979). In any event, popularization of the concept, although not the term, is evident in Burt and Dulay's report (1980) on their findings for Hispanic children in selected districts in California. It is ironic that whereas Skutnabb-Kangas and Toukomaa used their findings to support home language instruction programs through the early grades, Burt and Dulay imply that Hispanic CLs should be treated as L1 English speakers (although given the Lau procedures we know their home language is reported as Spanish). While it may be the case that some of the Hispanics diagnosed as limited proficiency in English and formerly

assumed to be fluent in Spanish are in fact English-superior, and not in need of bilingual education any more than monolingual LEPs, the CL category cuts both ways. It is just as possible that a particular CL is Spanish-superior even though she/he tested as limited in Spanish, and that she/he would benefit more from bilingual education than from English immersion.

B) Problems in test administration. The situation in which a test for language proficiency is administered presents a number of areas in which problems which interfere with accurate identification of language proficiency can be located. Two problem areas immediately suggest themselves, of which the first two present the most serious challenges to solution:

1. Resistance/unfamiliarity on the part of the child.
2. Mechanical and other errors on the part of the tester.

Problem area #1: Resistance/unfamiliarity on the part of the child. Much sociolinguistic and ethnolinguistic research has demonstrated the important effect that social situation has on speech, not only in the kind of language which is produced on different occasions but also on the volume of language. With particular reference to the "verbal deprivation" issue, Labov (1969) demonstrated that change in situation can have a decisive effect on whether a child may exhibit monosyllabic, limited proficient behavior, or fluent, syntactically rich proficient behavior. Children confronted with a situation which may be interpreted as threatening by them e.g., a formal test situation in school, may adopt the strategy of saying as little as possible in order to protect themselves from imagined self-incrimination. To the extent that fluency measured in words (or content morphemes) per sentence is used to judge language proficiency, as it is, for example, on the BINL (Basic Inventory of Natural Language), underestimation of language proficiency may be the result of such defensive behavior. It is expected that a review of the

literature on test anxiety for bilinguals and other minority populations not accustomed to the types of testing situations and intentions in these situations will further elucidate this problem (cf. Kenneth Hill, Institute for Child Development at U. of Illinois). In attempting to solve this problem, information is needed on the situational contexts that maximize speech for such children.

Problem area #2: Mechanical and other errors on the part of the tester. Mechanical errors include unreliable transcription, coding and calculation of scores by the tester. It is not clear how much error of this type there is. However, an examination of test instruments like the BINL and BSM (Bilingual Syntax Measure) reveal ample opportunity for such error. Those who have worked with transcription of actual speech production, which the above-mentioned instruments use as data, are aware of the difficulties and painstaking efforts required for accurate transcription. A large source of error may be expected in transcribing what is actually said. Another source of error can be found in the next stage of coding (assigning a numerical value to) what is said. In some cases this involves more than a simple count; the coder must recognize certain syntactic structures which are to be assigned a point value. Analysis of syntactic structures from speech is often problematic to linguists and would be of much greater difficulty to non-linguists. Finally, actual calculation of language proficiency scores from completed tests is a point for possible error, including gross error, and no doubt also a source of irritation to many reluctant testers. In point of fact, the dimensions of this problem have been such that ChecPoint Systems, Inc., the producers of the BINL test, have initiated a program of computerized calculation of language proficiency scores from transcription in order to accomodate the complaints of consumers of the test. Since there will still be some margin of error according to this process, and syntactic structures will have to be parsed for point assignment in calculating the scores, we are interested in finding out more about how this is being done.

Other sources of error, which anticipate the following section on test design, include misunderstanding of nonstandard vernacular Spanish responses by testers trained only in standard literary Spanish, resulting in dismissing a nonstandard response and requesting another response, or in inaccurate coding of the given response.

C) Problems in test design. Since the purpose of the language proficiency test instruments is to test language proficiency, defects in the design of the instrument will result in inaccurate conclusions. A general definition of language proficiency is a problem. To begin with, it is not at all clear that all forms of language proficiency are testable under formal conditions, since the normal everyday use of language occurs under conditions in which attention is not focused on language. Furthermore, in everyday situations, language is not being overtly tested but is being used to facilitate communication and social interaction. The recognition that only a small proportion of the language proficiencies a person may use in everyday life are directly testable may prevent the gross misunderstanding of the results of language proficiency testing. A narrow understanding of language proficiency may be sufficient for test purposes. The crucial purpose of language proficiency testing is to determine whether or not a child has sufficient proficiency in a particular language to be given meaningful educational instruction in that language. In addition to the question of the validity of the cutoff points used to evaluate proficiency, already discussed above and subject to study, there are three problem areas in test design that can be identified:

1. Irrelevant or wrong questions or tasks for assessment.
2. Built-in discriminatory factors against nonstandard language features.
3. Uncontrollable fluctuations in performance on the part of the testee, giving a false impression of stable competence.

Problem area #1: Irrelevant or wrong questions or tasks for assessment. To some extent there is overlap between 1 and 2 above, since a factor which biases testing against nonstandard language features can be argued to be an irrelevant or wrong question. However, clearly fitting into this first problem area is the problem of selecting which features, out of the universe of language or language-related proficiencies, should be tested in order to determine whether the tested language is sufficiently developed in the child to be the best medium of instruction in favoring school achievement. There is continuing controversy in the field of language proficiency testing concerning commensurability between test results on different aspects of language proficiency, e.g., phonological and grammatical knowledge. For example, the issue has been raised whether there is an interrelationship between results of discrete-point tests (particular linguistic features) and integrative tests (a battery of different linguistic proficiencies) (cf. Farhady, 1979).

Perhaps the most radical point of view is reflected in the work of Oller and Perkins (1980), who claim that all forms of language proficiency and a number of other proficiencies are commensurate. In a series of papers, Cummins (see, e.g., 1979, 1979b, 1980, 1980b) has contested this view and distinguished two types of language proficiency, one of which is related to scholastic success and the other of which is not. Cummins claims that most forms of language proficiency are irrelevant to school achievement. Basically, his argument is that almost all children are competent in interpersonal communicative skills, but much fewer children are successful in school. This calls into question whether tests for fluency which count such things as MLU (mean length of utterance)--a factor, for example, on the BINL test--are relevant to ability to use a language for school achievement. While Cummins' basic proposal seems reasonable, it does not necessarily follow that success in school is at all related to any kind of language proficiency. In personal communication with Cummins (August, 1980), this writer has found that he does not seem adverse to

conceding that what he has been calling CALP (Cognitive/Academic Language Proficiency), and claiming is directly related to school achievement, may not be a language proficiency at all, but rather something else not yet clearly defined. However, continued use of the term CALP may result in misunderstanding of just what this feature or combination of factors is. In addition, Cummins continually alludes in his work to the observations of Skutnabb-Kangas and Toukomaa, who in their discussion of "semilinguals" note that there is no correlation between the results of their "language proficiency" testing of the children of Finnish migrant workers in Sweden, which show the children to test lower than their monolingual Swedish and Finnish counterparts, and the observations of parents, teachers and the children themselves that the children are fluent in both languages.

Skutnabb-Kangas and Toukomas' conclusion that the apparent fluency and proficiency of these children is a "linguistic facade" and "superficial" is extremely dangerous, especially in a social science, since it dismisses as inappropriate any common sense notions about language proficiency available to a majority of observers, including most teachers. It leaves the determination of the important aspects of "language proficiency" in the hands of a few "experts". For example, if taken literally, this notion would invalidate the section of California's new Bilingual Education Improvement and Reform Act of 1980 which deals with procedures for investigating possible CL cases:

> (c) For those pupils identified as bieng of limited English proficiency, a further assessment shall be made to determine the pupil's primary language proficiency, including speaking, comprehension, reading, and writing to the extent assessment instruments are available . . .
>
> 14 If the assessment conducted pursuant to this subdivision indicates that the pupil has no proficiency in the primary language, further assessment of the pupil's primary language skills including consultation with the pupil's parents and guardians, the classroom teacher, the pupil, or others who are familiar with the pupil's language ability in various environments shall be conducted. . . (State of California 1979-80 A.B. 507 SEC.9)

While the research on CLs which is proposed in this paper will be much concerned with being relevant to how to proceed in the further assessment of such cases, it is felt that there are no grounds, despite the theoretical position of Skutnabb-Kangas and Toukomaa, and possibly Cummins, to preclude the procedures recommended beginning on line 14 of the above cited document. On the contrary, the requirement that such further assessment be done will be invaluable data for further elucidating the problem. It is also worth mentioning, as Leap (1980) observes, that in the writings of Skutnabb-Kangas and Toukomaa there are no examples of actual speech behavior of these so-called "semilinguals." We know nothing more about their language proficiencies than their collective test scores on certain test instruments. (See Dietrich & Freeman, 1979, for extended discussion of possible problems in assumptions underlying actual test items.)

Problem area #2: Built-in discriminatory factors against nonstandard language features. One most potent possibility for inadequate test design is linguistic prejudice against nonstandard varieties of speech being built into the test instrument (as well as into the test situation, as mentioned above under that section). If low test results are interpreted as meaning lack of proficiency in ANY variety of the tested language (rather than simply in the standard variety), as is often the case for prescriptively inclined educators who lack knowledge of nonstandard varieties of a language and/or are under the misconception that these are distorted or degenerate versions of the standard laguage, then the problem is in the test design. Tests which have these features are actually culturally unfair and biased because they dismiss language proficiencies which derive from the child's cultural background, in effect dismissing them as worthless. For example, the BSM, which dwells on the inflectional morphology in English (and other languages), gives points for many inflections which are characteristically absent or rare in the speech of many Black monolingual English speakers, e.g., third-person singular -s and the

possessive marker of standard English ('s). Many Black students taking the BSM test would score as LEP, although this would be culturally biased assessment of their language proficiency. Similarly, this defect applies to Hispanic bilinguals and appears to be based either on a lack of knowledge or oversight in the test design. For example, for the Spanish BSM the tester is cautioned not subtract points for omission of final -s in Spanish, since it is recognized that final -s deletion is a widespread dialect feature of many varieties of Spanish. Thus, the pronunciation of a word like las the feminine plural article, as homophonous with la may be a pronunciation feature rather than a sign of failure to have acquired the plural morpheme. However, English students are penalized for failure to pronounce the past tense suffix -ed despite copious documentation that when -ed is expected to occur immediately after a consonant, as in a word liked stopped, it is likely to be deleted for phonological rather than grammatical reasons in most varieties of spoken English.

There is possibility that any test design may be inadequate because it is addressed to the wrong language or variety of a language, causing the false impression that the speaker is actually deficient in any language. With respect to this, some scholars have raised the notion that in some areas a "mixed" language historically deriving from Spanish and English may have arisen so that one could not classify the language as either Spanish or English. This is distinct from the phenomenon of code-switching discussed previously, since code-switching implies that the two languages are separate at least on some occasions, although they may alternate even within the same sentence on others. De Avila (1979) following Laosa (1975) discusses this possibility referring to the mixed language as "Spanglish." It remains very much to be seen whether such a language exists as a unified code. Very little research on the subject exists. One of the few studies which directly addresses the problem is Huerta (1977). She studies a two-year old boy (until 2:10) from a middle-class background, who is primarily exposed to Spanish/English

code-switching from members of his family rather than Spanish and English separately. She demonstrates that the boy acquires both Spanish and English vocabulary and realizes that there are synonymous relations between pairs of words, e.g., "orange" and "naranja", "coffee" and "cafe" etc. She also claims that his syntactic development in both languages is comparable to what is known about monolingual speakers of either language at the same age. In concluding, she claims that despite these observations the child shows no evidence of recognizing that there are two languages involved. It remains for further research to establish whether a single unified language can evolve from extensive exposure to code-switching and whether such a unified language can be maintained in any social environment in the U.S. There is no evidence in Huerta's study to decide whether the child is dealing with one language or two from his own perspective. As for "Spanglish," the term has taken on condescending and derogatory connotations for many people and is used alternately to mean either code-switching, in which case two languages are involved, or the variety of Spanish spoken in many Mexican-American communities, in which case a nonstandard form of Spanish is involved with obvious influence of English historically (but not necessarily on the level of the individual speaker). In either case, from the perspective of the standard language, both code-switching and nonstandard Spanish tend to be overtly stigmatized and misinterpreted by the unwitting as a verbal deficit.

In the proposed research it will be important to study to what extent changes of situation change the proportion of the Spanish and English in conversation so that it can be determined whether or not Spanish and English act independently of each other and, thus, belong to separate language systems for individual users. It is interesting to note that in the Texas longitudinal study the alternating use of Spanish and English is sensitive to situation for the young school-age children, so that alternating language is most favored in the playground and least in the classroom. In addition, most switching

appears to be either at the sentence level or of single nominals (identified as "lexical") within the sentence. Since these children appear to be much stronger in Spanish than in English, as a whole, and close to the Mexican border, the situation might change considerably at an older age and in an area more distant from the border, such as Los Angeles, New York City or Chicago.

Problem area #3: Uncontrollable fluctuations in performance on the part of the testee, giving false impression of stable competence. Uncontrollable fluctuations in performance are real factors that interfere with getting an accurate assessment of a constant or base proficiency in any one testing period. Test designers have been aware of this problem and have cautioned testers against submitting children to test situations when they are ill or in some other visible way not in their "normal condition." Certainly the problem of ascertaining when a child is in the best condition to take a test is a problem of test administration and not test design. However, some fluctuations still little understood may interfere with consistent performance on a test from one day to the next and may be inherent in the test design.

Issues and Questions

Considered as a whole, this section has presented a formidable array of possible problems in obtaining accurate results from available language proficiency assessment instruments. It is essential that research and analysis be addressed to all of these possibilities, and that evaluation be made of which of these problems are the most prevalent and most urgently in need of correction.

To add substance to this theoretical section a few of the types of cases that have come to my attention are here presented.

MA, 8, a third-grade female born in Los Angeles, has tested as non-proficient in both English and Spanish. Does she need special

education because of a biologically based impairment or a serious psychological defect? Did she simply refuse to talk? Although her home language was reported as Spanish, is there another language in which she is proficient?

JR, 8, a third-grade male born in Mexico, tested as limited in English and non-proficient in Spanish. He has been in the LA school system since kindergarten. Is he actually more proficient in English than in Spanish?

RV, 10, a fifth-grade male born in Mexico and entering the LA school system the previous year. He tested as a non-speaker of English, as might expected, but also as a limited speaker of Spanish with a score of 38.6 on the BINL. What is his Spanish like? What happened during the test situation? How did he view it?

FR, 10, a fifth-grade male born in Los Angeles and in the same classroom as RV above. He tested as limited in both English and Spanish. Which is his stronger language? His Spanish score of 74.3 is considerably higher than his classmate RV's. Is he actually more proficient than RV in Spanish or is he simply more familiar with the test situation?

PS, 10, a fifth-grade female born in Los Angeles. In 1978 she tested as proficient in English according to the BINL. Inadvertently she was tested again in 1980, although there is no further procedure for proficiency testing once a child scores proficient. In 1980 she tested as a <u>limited English speaker</u>. This case was an accident, but either she was given someone else's score initially, or for some reason she did not perform at capacity on the retesting.

There are many other children that can be represented here. One thing we want to know is: How many? How many different types are

there? What is the reality underlying these test results? What language should they receive their primary instruction in?

Summary

1. The size and extent of the group labeled Assessed/Alleged Comparably Limited (ACL or "semilingual") is not well documented.

2. Legislators and educators are in need of research information on ACL pupils in order to formulate an equitable educational policy.

3. The results of language proficiency assessment instruments are open to question for accuracy and validity on a variety of bases, of which the most salient are a) norming calibration (what population do they rely on to norm language proficiency instruments?); b) equivalency (the extent to which different instruments agree in their assessment of any individual or group of children); c) intrinsic design (what they measure and how they score); and d) administration (effect of situation on performance).

4. The notion of language proficiency has a common sense base but is not clearly developed.

5. The actual speech behavior of the ACL and other children from bilingual and non-English speaking communities in the U.S. (especially between grades 3 and 6, ages 8 through 12) has not been systematically studied.

Hypotheses

1. Except for extremely high scorers on most language proficiency instruments, the instruments are not an accurate indicator of actual language proficiency or actual and potential school achievement.

2. Naturalistic discourse and interactional studies of actual speech behavior will provide a more accurate measure of actual language proficiency and will be more useful as a diagnostic tool.

3. The ACL category is largely an artifact of inaccurate instrumental measurements. It is the result of cultural and linguistic bias in testing.

4. School achievement does not correspond in any simple or direct way to either tested or actual language proficiency. That is, whether or not language is also involved, there are invariably other factors which are partially responsible for school achievement.

REFERENCES

Adams, Marianne L., 1980. Reference papers compiled for testing kit workshop III. School of Language Studies. Foreign Service Institute.

Attinasi, John et al., 1978. Language policy in the Puerto Rican community. Bilingual Review 5:1-2.

Berdan, R. H., 1979. National Advisory Council on Bilingual Education, 1978-1979. Fourth Annual Report, InterAmerica Research Associates, Inc. Rosslyn, Va.

Bickerton, Derek & C. Odo. 1976-1977. Study of Varieties of English in Hawaii, Vols. I & II. University of Hawaii.

Brent-Palmer, Cora, 1979. A sociolinguistic assessment of the notion 'im/migrant semilingualism' from a social conflict perspective. Working Papers on Bilingualism. 17. April, 1979. pp 135-180.

Burt, Marina & Heidi Dulay, 1976. Technical Manual for the Bilingual Syntax Measure.

Burt, Marina & Heidi Dulay, 1980. Relative proficiency of limited English-proficient students. Paper presented at 31st Annual Georgetown University Round Table on Languages and Linguistics.

Burt, Marina & Heidi Dulay, 1980b. "Hispanic students may be misdiagnosed." California Teacher, April, 1980.

California State Assembly Bill 507. 1980. Bilingual Education Improvement and Reform Act of 1980.

Cazden, B., V.P. John & D. Hymes, eds., 1972. Functions of Language in the Classroom. Teachers College Press; Columbia University, New York.

Cohen, Andrew D., Margret Bruck & Flora V. Rodriguez-Brown. 1979. Evaluating Evaluation: Bilingual Education Series 6. Papers in Applied Linguistics. Center for Applied Linguistics.

Cummins, James, 1979. Linguistic interdependence and the educational development of bilingual children. Review of Educational Research, 49:2, 222-51.

Cummins, James. 1979b. Cognitive/Academic Language Proficiency, Linguistic Interdependence, the optimum age question and some other matters. Working Papers in Bilingualism. 19. November, 1979. 197-205.

Cummins, James, 1980. The construct of language proficiency in bilingual education. Paper presented at the 31st Annual Georgetown University Round table on Languages and Linguistics, Georgetown University. March, 1980.

Cummins, James, 1980b. The cross-lingual dimensions of language proficiency: implications for bilingual education and the optimal age issue. TESOL Quarterly 14:2. 175-187.

De Avila, E.A. and Sharon E. Duncan, 1977. Language Assessment Scales. LAS I for Grades K-5. Linguametrics Group, Inc. Rio Piedras.

De Avila, Edward and Sharon Duncan, 1979. Relative linguistic proficiency and field dependence/independence: Some findings on linguistic heterogeneity and cognitive style of bilingual children. The Division of Bilingual and International Education. Domingo Dominguez, director. Southwest Educational Development Laboratory. Austin, Texas. NIE G-78-0208.

De Avila, E., S.E. Duncan, D. M. Ulibarri & J.S. Fleming, 1979. Predicting the success of language minority students from developmental, cognitive style, linguistic, and teacher perception measures. SEDL. NIE G-78-0208.

De Avila, E., & S.E. Duncan, 1979. Cheap Ship Trips: a preliminary study of some English phonological difficulties of languages minority students. SEDL. NIE-G-78-0208.

de Coster, Sylvian & Emile Derume, 1962. Retard pedagogique et situation sociale dans la region du Centre et au Borinage. Universite libre de Bruxelles, Institute de sociologie.

Dietrich, Thomas G. & C. Freeman, 1979. A linguistic guide to English proficiency testing in schools. Language in Education: Theory and Practice 23. Center for Applied Linguistics: December, 1979.

Dietrich, Thomas, G., C. Freeman & J-A. Crandall, 1979. A linguistic analysis of some English proficiency tests. Paper presented at the NABE conference. Seattle, May, 1979.

Epstein, N., 1977. Language, Ethnicity and the Schools: Policy Alternative for Bilingual-Bicultural Education. Washington D.C.: The George Washington University Institute for Educational Leadership.

Espino, Linda & Domingo Dominguez, 1980. Patterns of language mixing among Mexican American children in South Texas. A longitudinal study of the oral language development of Texas bilingual children (Spanish-English): Findings from the first year. Southwest Educational Development Laboratory.

Farhady, Hossein, 1979. The disjunctive fallacy between discrete-point and integrative tests. TESOL Quarterly 13:3., 347-357.

Federal Register. Department of Education. Vol. 45 No. 152 Tuesday, August 5, 1980. Proposed Rules. Problems of students with limited English proficiency.

Galang, Rosita, 1980. Language proficiency assessment of Asian-American populations. Paper presented at the Asian-Pacific Research Workshop, NCBR. Los Alamitos, September 4, 1980.

Garcia, Mary Ellen, 1980. Linguistic proficiency: How bilingual discourse can show that a child has it. Paper presented at the Ethnoperspectives in Bilingual Educational Research Forum. Eastern Michigan University, Ypsilanti, MI, June 12, 1980.

Gillmore, Gerry & Alice Dickerson, 1979. The relationship between instruments used for identifying children of limited English speaking ability in Texas. Region IV Education Service Center. Houston, Texas.

Gonzales, Gustavo, 1977. Teaching bilingual children. Bilingual Education: Current Perspectives, Vol. 2, Linguistics. Center for Applied Linguistics, Arlington, Va.

Guthrie, L. & M. S. Steffersen, 1980. Effect of situation on the verbalization of Black inner-city children. Paper presented at the Summer LSA meeting. Albuquerque, August 3, 1980.

Hansegard, N. E., 1975. Tvasprakighet eller halvsparkighet? (Bilingualism or Semilingualism?). Aldusserien 253, Stockholm.

Hernandez-Chavez, E, A. D. Cohen & A. Bellramo, 1975, Eds. El Lenguaje de los Chicanos: Regional & Social Characteristics of the language of Mexican-Americans. C.A.L., Arlington, Va.

Huerta, Ana, 1977. The acquisition of bilingualism: A code-switching approach. Sociolinguistic Working Paper 39. Southwest Educational Development Laboratory. Austin.

Jackson, Stephan L., 1980. Analysis procedures and summary statistics of the language data. A longitudinal study of the oral language development of Texas bilingual children (Spanish-English): Findings from the first year. Southwest Educational Development Laboratory.

Labov, William, 1968. Study of the New-Standard English of Negro & Puerto Rican Spoken in New York City. Final Report. NSF-3288.

Labov, William, 1972. The logic of nonstandard English. Language in the Inner City: Studies in the Black English Vernacular. University of Pennsylvania Press; Philadelphia, Pa.

Loasa, Luis M., 1975. Nonbiassed assessment of children's abilities: Historical antecedents and current issues. In R.L. Jones & B. Spolsky, eds. Testing Language Proficiency. Center for Applied Linguistics. Arlington, Va. 1-20.

Loasa, Luis M., 1975. Bilingualism in three U.S. Hispanic groups. Journal of Educational Psychology, 5.

Lambert, Wallace E., 1975. Culture and language as factors in learning and education. In Aaron Wolfgang, ed. Education of Immigrant Students: Issues and Answers. Toronto: Ontario Institute for Studies in Education. 55-83.

Lavandera, Beatriz, 1978. The variable component in bilingual performance. Paper presented at the 29th Annual Georgetown University Round Table on Languages and Linguistics: International Dimensions of Bilingual Education. March 17, 1978.

Leap, William L., 1979. On the implicit politics of educational research: Indian English and the "semilingualism conspiracy." Unpublished paper.

Linguametrics Group, 1977. Summary Report on Oral Language Assessment in Three Elementary Schools. Research Report #77-6.

Locks, Nancy A., B. Pletcher & D.F. Reynolds, 1978. Language Assessment Instruments for Limited-English-Speaking Students. A Needs Analysis. U.S. Dept. of HEW. NIE. Ocotober, 1978.

Mace-Matluck, Betty J., 1980. General characteristics of the children's language use in three environments. A Longitudinal Study of the Oral Language Development of Texas Bilingual Children (Spanish-English): Findings from the First Year. Southwest Regional Development Laboratory.

Mace-Matluck, Betty J. & J. H. Matluck, 1980. Oral language skills and reading performance. Southwest Educational Development Laboratory.

Matluck, Joseph H., 1980. An analysis of the children's phonological and grammatical structures. A longitudinal study of the oral language development of Texas bilingual children (Spanish-English): Findings from the first year. Southwest Educational Development Laboratory.

Macnamara, John, 1969. How can one measure the extent of a person's bilingual proficiency? In L.G. Kelly, ed. Description and Measurement of Bilingualism: an international seminar. University of Moncton, June 6-14, 1967. University of Toronto Press, pp 79-120.

Ministry of Education, 1953. The place of Welsh and English in the schools of Wales. Central Advisory Council for Education. London, Her Majesty's Stationery Office.

Oller, John W. & K. Perkins, 1980. Research in Language Testing. Newbury House Publishers, Inc.; Rowley, Mass.

Paulston, Christina Bratt, 1977. Research. Bilingual Education: Current Perspectives, Vol. 2, Linguistics. Center for Applied Linguistics, Arlington, Va.

Perlman, Carole L. and William K. Rice, Jr., 1979. A normative study of a test of English language proficiency. Chicago Board of Education. Department of Research and Evaluation.

Poplack, Shana, 1979. Sometimes I'll start a sentence in Spanish y termino en espanol: Toward a typology of code-switching. Centro Working Paper 4, Centro de estudios puertorriquenos, City University of New York; New York.

Rosansky, Ellen J., 1978. A Review of the Bilingual Syntax Measure. B. Spolsky, Ed. Advances in Language Testing. Series 1. Papers in Applied Linguistics. C.A.L. Arlington, VA.

Skutnabb-Kangas, T. & P. Toukomaa, 1976. "Teaching migrant children's mother tongue and learning the language of the host country in the context of the sociocultural situation of the migrant family." Department of Sociology and Social Psychology, University of Tampere, Finland.

Skutnabb-Kangas, T. & P. Toukomaa, 1979. Semilingualism and middle-class bias: A reply to Cora Brent-Palmer. Working Papers on Bilingualism. 19. November, 1979. pp 181-196.

Toukomaa, P. & T. Skutnabb-Kangas, 1975. The intensive teaching of the mother tongue to migrant children at preschool age. Department of Sociology and Social Psychology. University of Tampere, Finland.

Troike, Rudolph C., 1978. Research evidence for the effectiveness of bilingual education. National Clearinghouse for Bilingual Education. Arlington, VA.

Vildomec, Veroboj, 1963. Multilingualism. A.W. Sythoff; Leyden.

Wald, Benji, 1979. Limitations on the variable rule applied to bilingual phonology: The unmerging of the voiceless palatals in the English of Mexican American in the Los Angeles area. David Sankoff, ed. Recherches Linguistiques a Montreal. Universite de Quebec a Montreal.

Wald, Benji, 1974. Bilingualism. Annual Review of Anthropology Vol. 3, 301-21.

Weinreich, Uriel, 1967. Languages in Contact. Findings and Problems. Mouton & Co., the Hague.

Wolfram, Walt, D. Christian, W. L. Leap & L. Potter. Variability in the English of Two Indian Communities and its Effect on Reading & Writing. Final Report. NIE-G-77-0006. Center for Applied Linguistics, Arlington, VA.